METHODS AND COMPOSITIONS FOR RNA-DIRECTED TARGET DNA MODIFICATION AND FOR RNA-DIRECTED MODULATION OF TRANSCRIPTION

CROSS-REFERENCE

This application is a continuation of U.S. patent application Ser. No. 13/842,859, filed Mar. 15, 2013, which claims the benefit of U.S. Provisional Patent Application Nos. 61/652,086 filed May 25, 2012, 61/716,256 filed Oct. 19, 2012, 61/757,640 filed Jan. 28, 2013, and 61/765,576, filed Feb. 15, 2013, each of which applications is incorporated herein by reference in its entirety.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING PROVIDED AS A TEXT FILE

A Sequence Listing is provided herewith as a text file, "BERK-187-SeqList_ST25.txt" created on Mar. 14, 2013 and having a size of 7645 KB. The contents of the text file are incorporated by reference herein in their entirety.

BACKGROUND

About 60% of bacteria and 90% of archaea possess CRISPR (clustered regularly interspaced short palindromic repeats)/CRISPR-associated (Cas) system systems to confer resistance to foreign DNA elements. Type II CRISPR system from *Streptococcus pyogenes* involves only a single gene encoding the Cas9 protein and two RNAs—a mature CRISPR RNA (crRNA) and a partially complementary trans-acting RNA (tracrRNA)—which are necessary and sufficient for RNA-guided silencing of foreign DNAs.

In recent years, engineered nuclease enzymes designed to target specific DNA sequences have attracted considerable attention as powerful tools for the genetic manipulation of cells and whole organisms, allowing targeted gene deletion, replacement and repair, as well as the insertion of exogenous sequences (transgenes) into the genome. Two major technologies for engineering site-specific DNA nucleases have emerged, both of which are based on the construction of chimeric endonuclease enzymes in which a sequence non-specific DNA endonuclease domain is fused to an engineered DNA binding domain. However, targeting each new genomic locus requires the design of a novel nuclease enzyme, making these approaches both time consuming and costly. In addition, both technologies suffer from limited precision, which can lead to unpredictable off-target effects.

The systematic interrogation of genomes and genetic reprogramming of cells involves targeting sets of genes for expression or repression. Currently the most common approach for targeting arbitrary genes for regulation is to use RNA interference (RNAi). This approach has limitations. For example, RNAi can exhibit significant off-target effects and toxicity.

There is need in the field for a technology that allows precise targeting of nuclease activity (or other protein activities) to distinct locations within a target DNA in a manner that does not require the design of a new protein for each new target sequence. In addition, there is a need in the art for methods of controlling gene expression with minimal off-target effects.

SUMMARY

The present disclosure provides a DNA-targeting RNA that comprises a targeting sequence and, together with a modifying polypeptide, provides for site-specific modification of a target DNA and/or a polypeptide associated with the target DNA. The present disclosure further provides site-specific modifying polypeptides. The present disclosure further provides methods of site-specific modification of a target DNA and/or a polypeptide associated with the target DNA The present disclosure provides methods of modulating transcription of a target nucleic acid in a target cell, generally involving contacting the target nucleic acid with an enzymatically inactive Cas9 polypeptide and a DNA-targeting RNA. Kits and compositions for carrying out the methods are also provided. The present disclosure provides genetically modified cells that produce Cas9; and Cas9 transgenic non-human multicellular organisms.

Features

Features of the present disclosure include a DNA-targeting RNA comprising: (i) a first segment comprising a nucleotide sequence that is complementary to a sequence in a target DNA; and (ii) a second segment that interacts with a site-directed modifying polypeptide. In some cases, the first segment comprises 8 nucleotides that have 100% complementarity to a sequence in the target DNA. In some cases, the second segment comprises a nucleotide sequence with at least 60% identity over a stretch of at least 8 contiguous nucleotides to any one of the nucleotide sequences set forth in SEQ ID NOs:431-682 (e.g., 431-562). In some cases, the second segment comprises a nucleotide sequence with at least 60% identity over a stretch of at least 8 contiguous nucleotides to any one of the nucleotide sequences set forth in SEQ ID NOs:563-682. In some cases, the site-directed modifying polypeptide comprises an amino acid sequence having at least about 75% amino acid sequence identity to amino acids 7-166 or 731-1003 of the Cas9/Csn1 amino acid sequence depicted in FIG. **3**A and FIG. **3**B, or to the corresponding portions in any of the amino acid sequences set forth as SEQ ID NOs: 1-256 and 795-1346.

Features of the present disclosure include a DNA polynucleotide comprising a nucleotide sequence that encodes the DNA-targeting RNA. In some cases, a recombinant expression vector comprises the DNA polynucleotide. In some cases, the nucleotide sequence encoding the DNA-targeting RNA is operably linked to a promoter. In some cases, the promoter is an inducible promoter. In some cases, the nucleotide sequence encoding the DNA-targeting RNA further comprises a multiple cloning site. Features of the present disclosure include an in vitro genetically modified host cell comprising the DNA polynucleotide.

Features of the present disclosure include a recombinant expression vector comprising: (i) a nucleotide sequence encoding a DNA-targeting RNA, wherein the DNA-targeting RNA comprises: (a) a first segment comprising a nucleotide sequence that is complementary to a sequence in a target DNA; and (b) a second segment that interacts with a site-directed modifying polypeptide; and (ii) a nucleotide sequence encoding the site-directed modifying polypeptide comprising: (a) an RNA-binding portion that interacts with the DNA-targeting RNA; and (b) an activity portion that exhibits site-directed enzymatic activity, wherein the site of enzymatic activity is determined by the DNA-targeting RNA.

Features of the present disclosure include a recombinant expression vector comprising: (i) a nucleotide sequence encoding a DNA-targeting RNA, wherein the DNA-targeting RNA comprises: (a) a first segment comprising a nucleo-